ASYMPTOTIC DISTRIBUTION OF THE SHAPIRO-WILK W FOR TESTING FOR NORMALITY

BY

J. R. LESLIE, M. A. STEPHENS, and S. FOTOPOULOS

TECHNICAL REPORT NO. 411

DECEMBER 22, 1988

PREPARED UNDER CONTRACT
N00014-86-K-0156 (NR-042-267)
FOR THE OFFICE OF NAVAL RESEARCH

DEPARTMENT OF STATISTICS
STANFORD UNIVERSITY
STANFORD, CALIFORNIA

AD-A202 269

ASYMPTOTIC DISTRIBUTION OF THE SHAPIRO-WILK W FOR TESTING FOR NORMALITY

BY

J. R. LESLIE, M. A. STEPHENS, and S. FOTOPOULOS

TECHNICAL REPORT NO. 411
DECEMBER 22, 1988

Prepared Under Contract
N00014-86-K-0156 (NR-042-267)
For the Office of Naval Research

Herbert Solomon, Project Director

DEPARTMENT OF STATISTICS
STANFORD UNIVERSITY
STANFORD, CALIFORNIA

ASYMPTOTIC DISTRIBUTION OF THE SHAPIRO-WILK W FOR TESTING FOR NORMALITY

BY

J. R. LESLIE, M. A. STEPHENS, and S. FOTOPOULOS

TECHNICAL REPORT NO. 411
DECEMBER 22, 1988

Prepared Under Contract
N00014-86-K-0156 (NR-042-267)
For the Office of Naval Research

Herbert Solomon, Project Director

DEPARTMENT OF STATISTICS
STANFORD UNIVERSITY
STANFORD, CALIFORNIA

ASYMPTOTIC DISTRIBUTION OF THE SHAPIRO-WILK W FOR TESTING FOR NORMALITY

By

J. R. Leslie, M. A. Stephens and S. Fotopoulos

1. Introduction.

A popular test for the normality of a random sample is based on the Shapiro-Wilk statistic W. This statistic, which was presented in Shapiro and Wilk (1965), is the ratio of the square of the BLUE of σ to the sample variance, where σ^2 is the variance of the normal population from which the sample is assumed, under the null hypothesis, to have been drawn. For convenience we shall work with $W^{1/2}$ which has the form

$$W^{1/2} = \underline{X}'V_0^{-1}\underline{m}/\{\sum_1^n (X_i-\bar{X})^2\underline{m}'V_0^{-1}V_0^{-1}\underline{m}\}^{1/2} ,$$

where $\underline{X} = (X_1,\ldots,X_n)'$, $X_1 < X_2 < \ldots < X_n$, is the vector of order statistics from the sample, $\bar{X}$ is the sample mean, and $\underline{m}$ is the mean vector and V_0 the covariance matrix of standard normal order statistics. As $W^{1/2}$ is location and scale invariant we can assume from henceforth that $X_1,\ldots,X_n$, are order statistics for a sample from a $N(0,1)$ population.

A number of authors (for example, Sarkadi (1975), (1977) and Gregory (1977)) have (correctly) guessed at the form of the asymptotic distribution for W as well as predicting that the test should be consistent. However no rigorous proofs have been possible due to the presence of V_0^{-1}. Neither V_0 nor V_0^{-1} can be found explicitly and until recently no

reasonably accurate asymptotic approximation for V_0 was available. A paper by one of the authors (Leslie 1984) has now remedied the situation; in that paper can be found an approximation for V_0 together with a number of asymptotic properties of V_0, one of which is of particular importance to this work. It states that $\underline{m}$ is approximately an eigenvector of V_0^{-1} in the following sense:

$$(1) \qquad // V_0^{-1}\underline{m} - 2\underline{m} // \leq C(\log n)^{-1/2} ,$$

where C is a constant independent of n, and $//\underline{b}//^2 = \Sigma b_i^2$ for $\underline{b} = (b_1,\ldots,b_n)'$. This latter result formalises a similar one appearing in Stephens (1975).

The asymptotic distribution of W, after appropriate normalizing, has been assumed to be the same as that of the De Wet and Venter (1972) statistic

$$W^* = r^2(\underline{X},\underline{H}) ;$$

here $r(\underline{X},\underline{Y})$ is the sample correlation coefficient between $\underline{X}$ and $\underline{Y}$, $\underline{H}$ is the $n \times 1$ vector whose i^{th} element is $\Phi^{-1}\{i/(n+1)\}$ and $\Phi^{-1}(\cdot)$ is the inverse function for the standard normal distirbution function $\Phi(\cdot)$, that is $\Phi^{-1}(\Phi(x)) = x$.

The rationale behind this assumption was that firstly, $V_0^{-1}\underline{m}$ was known to behave like $2\underline{m}$ (see Stephens (1975)), secondly, $\Phi^{-1}\{i/(n+1)\}$ approximates the i^{th} element of $\underline{m}$ and thirdly, as V_0 is a doubly stochastic matrix (the sum along any row or column is 1) we may write

$$W = r^2(\underline{X}, V_0^{-1}\underline{m}) \ .$$

De Wet and Venter (1972) showed that the asymptotic distribution of W^* has the form

$$2n(1-W^{*1/2}) - a_n \xrightarrow{D} \zeta \tag{2}$$

where $\zeta = \Sigma_3^\infty (Y_i^2-1)/i$, $\{Y_i, i \geqq 1\}$ is a sequence of i.i.d. $N(0,1)$ variates,

$$a_n = (n+1)^{-1} \{ \sum_{i=1}^{n} j(1-j)(\phi\{\Phi^{-1}(j)\})^{-2} - \frac{3}{2} \tag{3}$$

$j = i/(n+1)$ and $\phi(\cdot)$ is the $N(0,1)$ density function.

Beyond the De Wet and Venter result the first step towards the asymptotic distribution for W was to show that the Shapiro-Francia (1972) statistic $W^\dagger$ given by

$$W^\dagger = r^2(\underline{X}, \underline{m}),$$

behaves in the same way as W^*. This was done independently and via different routes by Verrill and Johnson (1983) and by the authors in Fotopoulos, Leslie, and Stephens (1984), henceforth called FLS, where expression (2) was established with $W^\dagger$ in place of W^*. In fact we show in FLS the equivalent result that

$$n(W^{*1/2} - W^{\dagger 1/2}) \to 0 \text{ in probability.} \tag{4}$$

Our task in the present paper is to show that

$$n(W^{1/2}-W^{\dagger 1/2}) \to 0 \quad \text{in probability} . \tag{5}$$

We note that Verrill and Johnson (1983) contains a result (Theorem 3) which should eventually cover the asymptotic distribution of W. However certain properties of $V_0^{-1}\underline{m}$ need to be established before it can be applied. Inequality (1) does not appear to be enough.

2. Asymptotic Properties of W and a_n.

The following theorem presents one version of the asymptotic distribution for W - in fact the asymptotic distribution for $W^{1/2}$ - whilst the corollary offers the complementary form in terms of W.

Theorem. Under the hypothesis that the observed sample is from a normal population the asymptotic distribution of the Shapiro-Wilk W takes the form:

$$2n(1-W^{1/2}) - 2En(1-W^{1/2}) \overset{D}{\to} \zeta$$

where $\zeta = \Sigma_3^\infty (Y_i^2-1)/i$, and $\{Y_i, i \geq 3\}$ is a sequence of i.i.d. $N(0,1)$ variables.

From the lemma below and from the theorem we have $\sqrt{n}(1-W^{1/2}) \to 0$ in probability, which leads to

$$2n(1-W^{1/2}) - n(1-W) = (\sqrt{n}(1-W^{1/2}))^2 \to 0 \quad \text{in probability.}$$

Again applying the lemma below we obtain,

Corollary. An equivalent form for the asymptotic distribution of W is:

$$n(W-EW) \xrightarrow{D} -\zeta \ .$$

It is not obvious from their definition just how the constants a_n will behave as n gets large. The following lemma should shed some light on this matter.

Lemma. The constants a_n defined in (3) have the following properties:

(i) $a_n - 2nE\{1-r(\underline{X},\underline{b})\} \to 0$, where $\underline{b}$ can be any of $\underline{m}$, $1/2V_0^{-1}\underline{m}$ or $\underline{H}$,

(ii) $a_n - nE(1-W) \to 0$,

(iii) $|a_n - n(1-n^{-1}\underline{m}'\underline{m}) + 3/2| \leqq C(\log n)^{-1}$,

and

(iv) $C_1 \log\log(n) < a_n < C_2 \log\log(n)$, $0 < C_1 < C_2 < \infty$.

Note that (iii) implies that

$$\underline{m}'\underline{m} = n - a_n - 3/2 + 0(1) \ .$$

As far as we are aware this property of $\underline{m}'\underline{m}$ has not appeared elsewhere; the behavior of $\underline{m}'\underline{m}$ is of interest in other contexts and has been the subject of a number of papers (see for example, Balakrishnan (1984), Ruben (1956) and Saw and Chow (1966)).

It should be pointed out that the covergence for (i) and (ii) in the lemma is extremely slow; for example $a_n - 2En(1-r(\underline{X},\underline{m})) \simeq -0.1$ for $40 \leqq n \leqq 400$. It is therefore unclear as to which set of norming constants it is best to use.

When Sarkadi (1975) established the consistency of the Shapiro-Francia test, it seemed likely that the Shapiro-Wilk test would share that property. That it is indeed consistent will follow from a straightforward application of a result in Sarkadi (1981).

3. Proofs.

Notation. We give some notation which will be used throughout the rest of the paper. With or without subscripts, C is a generic constant which is independent of i and n. Set $\underline{g} = 1/2V_0^{-1}\underline{m}$, $nG_n^2 = \underline{g}'\underline{g}$, $nM_n^2 = \underline{m}'\underline{m}$, N is the integer part of $1/2(n+1)$, $S_n^2 = \Sigma_1^n (X_i-\overline{X})^2/n$, $\psi(v) = \Phi^{-1}\{\exp(-v)\}$, $s_i = \Sigma_n^i \nu^{-1}$, $\psi_i = \psi(s_i)$, $W_i = -\log(\Phi(X_i))-s_i$. Note W_i+s_i is the i^{th} largest order statistic in a random sample from an exponential distribution; $EW_i = 0$, $EW_i^2 = d_{in}$, where $d_{in} = \Sigma_i^n \nu^{-2}$, $EW_i^3 = 2\Sigma_i^n \nu^{-3}$ and $|EW_i^r| \leqq Ci^{-2}$ for $r \geqq 3$. Denote the i^{th} element of $\underline{g}$, $\underline{m}$ and $\underline{H}$ by respectively g_i, m_i and H_i ($\underline{m}$ and $\underline{H}$ are given in section 1). Further, as r is scale and location invariant we assume without loss of generality that our sample is from a $N(0,1)$ population.

Proof of Consistency. The consistency of W follows directly from Theorem 1 of Sarkadi (1981). There is a small difficulty in that whilst it appears to be the case that $V_0^{-1}\underline{m}$ is a vector whose elements, as you move down the vector, are monotonic increasing, we are unbale to prove it.

This means we cannot establish that $W^{1/2}$ is always positive. Sarkadi exploits the fact that $W^{+1/2}$ is always positive to argue that tests based on $W^{+1/2}$ are equivalent to those based on W^{+}. We need to argue likewise for W (note: we distinguish between $W^{1/2}$, $W^{+1/2}$ etc. and the square roots of W, W^{+} etc.; it is true that $W = (W^{1/2})^2$ but in view of what has just been said, we are unable to say whether $W^{1/2}$ is the positige square root of W). We overcome this difficulty by showing below that

$$(6) \qquad W^{1/2} \geqq -C(\log n)^{-1/2}, \quad C \text{ independent of } n.$$

From the theorem and the lemma, the 100α% critical region for the test based on $W^{1/2}$ is: $W^{1/2} < 1-1/2(c(\alpha)+a_n)n^{-1}$. For the test based on W it is $W < 1-(c(\alpha)+a_n)n^{-1}$. By (6) the two critical regions are asymptotically equivalent. We need only show therefore that $W^{1/2}$ is consistent. We extablish (6) by setting $\underline{1}_n$ to be an $n \times 1$ vector of 1's and writing

$$W^{1/2} = \{(\underline{X}-\bar{X}\underline{1}_n)'(\underline{g}-\underline{m}) + \underline{X}'\underline{m}\}/(nS_nG_n) .$$

As $\underline{X}'\underline{m} > 0$ provided only that the components of $\underline{X}$ are increasing (see Sarkadi (1975), Lemma 2), and from (1), $\max|g_i-m_i| < C/\sqrt{(\log n)}$, we have, with the help of (21) below,

$$W^{1/2} \geqq -C\Sigma_1^n|X_i\bar{X}|/\{nS_nG_n\sqrt{(\log n)}\} \geqq -C/\sqrt{(\log n)} .$$

We turn now to Theorem 1 of Sarkadi (1981). Applied to our context, it states that $W^{1/2}$ will determine a consistent test of H_0: that the random

sample is normal, versus H_1: that the observations are not normal (Sarkadi also allows the observations under H_1 to be m-dependent with common non-normal marginal) providing

$$(7) \qquad \sum_1^n g_i G_n^{-1} \int I(i-1 < nu < i)\Phi^{-1}(u)du = 1 + O(1)$$

where $I(A)$ is the indicator function for A. Note that Sarkadi's theorem is framed in terms of a statistic T_n which here takes the form

$$T_n = \sum_1^n \{(X_i-\bar{X})n^{-1/2}S_n^{-1} - c_{in}\}^2 = 2(1-W^{1/2}) \ ,$$

where $c_{in}\sqrt{n} = g_i/G_n$. To establish (7) we require results contained in the proof of both our lemma and theorem, therefore we will leave the derivation of (7) till the end of the article.

<u>Proof of Lemma</u>. We start by showing (iii); observe that

$$n(1-M_n^2) = 2\{\sum_1^N \mathrm{Var}(X_i)\} - (2N-n)\mathrm{Var}(X_N) \ .$$

We can write

$$\mathrm{Var}(X_i) = E\{\psi(s_i+W_i) - E\psi(s_i+W_i)\}^2 \ .$$

Expanding ψ in W_i up to third order terms, using the properties of W_i given in the section on notation and together with results in Leslie (1984) (in particular, Lemma 6 and the properties of ψ given in section 3) we can show that

$$|\mathrm{Var}(X_i) - \{\psi'(s_i)\}^2 d_{in}| \leqq C\{i(\log(n/i))\}^{-2} ,$$

where $\psi'(s_i) = \{\exp(-s_i)\}/\phi(\Phi^{-1}(\exp(-s_i)))$ and $d_{in} = \Sigma_i^n \nu^{-2}$. This yields

$$(8) \qquad |\sum_1^N \mathrm{Var}(X_i) - \sum_1^N \{\psi'(s_i)\}^2 d_{in}| \leqq C(\log n)^{-2} .$$

Using the Euler-Maclaurin summation formula (Knopp (1951), p. 534)

$$(9) \qquad 0 < s_i - \log((n+1)/i) - 1/2(i^{-1}-(n+1)^{-1}) < \{i^{-2}-(n+1)^{-2}\}/12$$

and

$$(10) \qquad 0 < d_{in} - i^{-1}\{1-(i/(n+1))\} - 1/2\{i^{-2}-(n+1)^{-2}\} < 1/(6i^3) .$$

In FLS we show that $|\psi'(v)|$ and $|\psi''(v)|$ are monotonic decreasing in v; also in Lemmas 1 and 4 in Leslie (1984) it is shown that

$$(11) \qquad |\psi'\{\log((n+1)/i)\}| < C\{\log(n/i)\}^{-1/2}$$

and

$$(12) \qquad |\psi''\{\log((n+1)/i)\}| < C\{\log(n/i)\}^{-3/2} .$$

With (9), (10) and (11) we have

$$(13) \qquad |\{\psi'(s_i)\}^2(d_n-i^{-1}\{1-(i/(n+1))\})| < Ci^{-2}/\log(n/i)$$

(14) $$\left|\{\psi'(s_i)\}^2 - (\psi'\{\log((n+1)/i)\})^2\right| < C\left|\psi'(\alpha_i)\psi''(\alpha_i)\right|/i ,$$

where $\log\{(n+1)/i\} < \alpha_i < s_i$.

Expressions (11)-(14) taken together imply that

(15) $$\left|\sum_1^N \{\psi'(s_i)\}^2 d_{in} - (n+1)^{-2} \sum_1^N (\phi\{\Phi^{-1}(i/(n+1))\})^{-2} i\{1-i(n+1)^{-1}\}\right|$$
$$< C(\log n)^{-1} .$$

From the definition of a_n and with (8) and (15) we obtain (iii).

Next we establish (iv). A well known inequality is useful here (see Renyi (1970) p. 164; for $x < 0$

(16) $$\phi(x)(1-x^{-2})/|x| < \Phi(x) < \phi(x)/|x| .$$

From this we obtain, for $1 \leqq i \leqq N$, and with $x = H_i$,

(17) $$1-H_i^{-2} \leqq i|H_i|/\{(n+1)\phi(H_i)\} \leqq 1 .$$

In view of the symmetry in the summands in a_n, we need consider only $1 \leqq i \leqq N$. We use (17), over the range $1 \leqq i \leqq [\frac{1}{2}N]$ and for $[\frac{1}{2}N] < i \leqq N$ we use

(18) $$C_1 < \phi(H_i)(i/(n+1))\{1-(i/(n+1))\} < C_2 ,$$

where C_1, C_2 do not depend on i or n. Based on (16) we show in Lemma 3

of FLS that for any $c_0 (0 < c_0 < 1)$ there is a $\gamma(c_0)$ such that when $0 < u < \gamma(c_0) < \frac{1}{2}$,

$$(19) \qquad -\{-\log(2\pi u^2)\}^{1/2} < \Phi^{-1}(u) < -\{-c_0 \log(2\pi u^2)\}^{1/2} .$$

This yields for $1 \leqq i \leqq N$,

$$(20) \qquad C_3\{\log(n/i)\}^{1/2} < |H_i| < C_4\{\log(n/i)\}^{1/2} .$$

Applying (17), (18) and (20) we find

$$C_5 \sum_{1}^{[\frac{1}{2}N]} \{i\log(n/i)\}^{-1} + C_6 < a_n + 3/2 < C_7 \sum_{1}^{[\frac{1}{2}N]} \{i\log(n/i)\}^{-1} + C_8$$

which, after approximating the sum by an integral, establishes (iv).

To complete the lemma we prove (i) and (ii). First however, we need two results which will be used here and in the proof of the theorem:

$$(21) \qquad G_n \to 1 \quad \text{as} \quad n \to \infty, \quad \text{and}$$

$$(22) \qquad 0 \leqq \|\underline{m}\|\,\|\underline{g}\| - \underline{m}'\underline{g} \leqq M_n G_n^{-1} \|\underline{g}-\underline{m}\|^2 .$$

It is well known that $M_n \to 1$ as $n \to \infty$ (see Hoeffding (1953)). On writing $G_n^2 = M_n^2 + 2\underline{m}'(\underline{g}-\underline{m})n^{-1} + \|\underline{g}-\underline{m}\|^2 n^{-1}$, from (1) and Schwarz inequality we obtain (21). We demonstrate (22) by exploiting an idea in Sarkadi (1972). First note that $\underline{m}'\underline{g} > 0$, for $\underline{m}'\underline{g} = \underline{m}'V_0^{-1}\underline{m}$ and V_0 being a covariance matrix, is positive definite. Set θ to be the angle

between $\underline{m}$ and $\underline{g}$ then $\cos\theta > 0$ and $0 < \theta < \frac{1}{2}\pi$. Consider the triangle formed by vectors $\underline{m}$, $\underline{g}$ and $\underline{a} = \underline{m}-\underline{g}$, respectively lines AB, AC and CB. Let CD be the perpendicular from C to AB. Then $\|\underline{a}\|^2 \geq (CD)^2 = \|g\|^2\sin^2\theta = \|\underline{g}\|^2(1+\cos\theta)(1-\cos\theta) > \|\underline{g}\|^2(1-\cos\theta) \geq 0$. As $\cos\theta = \underline{m}'\underline{g}\{\|\underline{m}\|\,\|\underline{g}\|\}$, (22) follows.

Returning to the proof of (i) and (ii) of the lemma we show first that

$$(23) \qquad |nE(1-r(\underline{X},\underline{b})) - n(1-M_n) + 3/4| \leq C(\log n)^{-1} .$$

As r is scale invariant and as S_n^2 is sufficient for the scale parameter σ we can use Theorem 7, p. 243 of Hogg and Craig (1970) to yield

$$nE\, r(\underline{X},\underline{b}) = \underline{m}'\underline{b}/\{ES_n \|\underline{b}\|\, n^{-1/2}\} .$$

With nS_n^2 distributed as χ^2 on $n-1$ degrees of freedom it is elementary to show that

$$ES_n = (2/n)^{1/2}\Gamma(n/2)/\Gamma((n-1)/2) .$$

By Stirling's formula this reduces to $1-(3/4)n^{-1} + O(n^{-2})$. As $n^{-1/2}\|\underline{b}\| \to 1$ (the case $\underline{b} = \underline{H}$ is shown in Lemma 2 of De Wet and Venter (1972)), and using (1), (22) and an analogue of (22) with $\underline{g}$ and G_n replaced by $\underline{H}$ and $H_n = \sqrt{\{(\underline{H}'\underline{H})/n\}}$ (this analogue holds because $m_iH_i > 0$ for all i, m_i and H_i always having the same sign) we have,

$$En\{1-r(\underline{X},\underline{b})\} = n(ES_n)^{-1}\{1-(3/4)n^{-1} - n^{-1/2}\underline{m}'\underline{b}/\!/\,\underline{b}/\!/^{-1}\} + 0(n^{-1})$$

$$= n(ES_n)^{-1}(1-M_n) - (3/4) + 0(\log n)^{-1} ,$$

$$= n(1-M_n) - (3/4) + 0(\log n)^{-1} ,$$

the latter expression resulting from the fact that $n(1-M_n) = 0(\log\log n)$ (using (iii) and (iv) of the lemma and recall that $M_n \to 1$). This establishes (23). Analogous to (23) for $\underline{b} = \underline{g}$ we have

$$(24) \qquad |nE\{1-r^2(\underline{X},\underline{g})\} - n(1-M_n^2) + 3/2| \leq C(\log n)^{-1} .$$

To show this we note that as $nES_n^2 = n-1$ we can write

$$nEr^2(\underline{X},\underline{g}) = E(\underline{X}'\underline{g})^2/\{(n-1)G_n^2\}$$

with

$$E(\underline{X}'\underline{g})^2 = \underline{g}'V_0\underline{g} + (\underline{g}'\underline{m})^2 = \tfrac{1}{2}\underline{m}'\underline{g} + (\underline{g}'\underline{m})^2$$

$$= \tfrac{1}{2}nM_nG_n + (n\,G_nM_n)^2 + 0(n/\log n)$$

using (1) and (22). Again using the property that $n(1-M_n) = 0(\log\log n)$, we obtain (24). As

$$(25) \qquad 2n(1-M_n) - n(1-M_n^2) = \{\sqrt{n}(1-M_n)\}^2 = 0\{(\log\log n)^2/n\}$$

it is clear from (23), (24) and (iii) of the lemma that (i) and (ii) hold.

Undoubtedly it is true that $a_n - E_n\{1-r^2(\underline{X},\underline{b})\} \to 0$, for $\underline{b} = \underline{m}$ and $\underline{H}$. However this entails showing that $/\!/ V_0\underline{H}-2\underline{H}/\!/ \to 0$ and $/\!/ V_0\underline{m}-2\underline{m}/\!/ \to 0$ both of which will follow once V_0 is replaced by the approximation V given in Leslie (1984): corollary 1 in Leslie (1984) permitting this. These two results will involve a quantity of tedious analysis and it seems unnecessary to set it down here.

<u>Proof of Theorem</u>. The theorem follows from (2), (4) and (5) together with the lemma; therefore to prove the theorem it remains to establish (5). Now

$$nS_n(W^{1/2}-W^{+1/2}) = \sum_1^n X_i(g_iG_n^{-1}-m_iM_n^{-1})$$

$$= \sum (X_i-m_i)(g_i-m_i)G_n^{-1} + \sum (X_i-m_i)m_i(G_n^{-1}-M_n^{-1}) + (\underline{m}'\underline{g} - /\!/ \underline{m}/\!/ \; /\!/ \underline{g}/\!/)G_n^{-1}.$$

As $S_n \to 1$ a.s. and with (21) and (22), expression (5) will follow from Markov's inequality once we demonstrate that

(26) $$E|\Sigma(X_i-m_i)(g_i-m_i)| \to 0, \quad \text{and}$$

(27) $$E|\Sigma(X_i-m_i)(G_n^{-1}-M_n^{-1})| \to 0 \; .$$

Result (26) follows from Schwarz inequality:

$$E|\Sigma(X_i-m_i)(g_i-m_i)| \leqq \{n(1-M_n^2)\}^{1/2} \; /\!/ \underline{g}-\underline{m}/\!/ \; .$$

With (1) and with (iii) and (iv) of the lemma we have (26).

To deal with (27) we note that in Lemma 11 of FLS we show

(28) $$E|X_i-\psi_i| < C/\sqrt{\{i\log(n/i)\}}$$

and in Theorem 1 in FLS we show that

(29) $$|\psi_i-m_i| < Ci^{-1}\{\log(n/i)\}^{-3/2} \ ;$$

both of these bounds hold provided $1 \leqq i \leqq N$. As $|//m//-//g//| \leqq //m-g//$,

(30) $$|G_n^{-1}-M_n^{-1}| \leqq //\underline{m}-\underline{g}// \,/(M_n G_n \sqrt{n}) \leqq C(n\log n)^{-1/2}$$

and

(31) $$E|\Sigma(X_i-m_i)m_i| \leqq 2\sum_1^N (E|X_i-\psi_i||m_i|+|\psi_i-m_i||m_i|) \ .$$

From (29), (9), (20) and the monotonicity (decreasing) of $|\psi(v)|$

(32) $$|m_i| \leqq C\{\log(n/i)\}^{1/2}, \quad 1 \leqq i \leqq N \ ,$$

so by combining results (28) to (32) we find

$$E|\Sigma(X_i-m_i)m_i\{G_n^{-1}-M_n^{-1}\}| \leqq C(\log n)^{-1/2} \ .$$

This establishes (27) and hence the Theorem.

Derivation of Expression (7). Denote the integral in (7) by $J(i,n)$ then

$$J(i,n) = \begin{cases} -\phi\{\Phi^{-1}(1/n)\}, & \text{for } i = 1 , \\ \Phi^{-1}((i-1)/n)n^{-1} + \frac{1}{2}n^{-2}(\phi\{\Phi^{-1}((i-\theta)/n)\})^{-1}, & 0 < \theta < 1,\ 1 < i < n \\ \phi\{\Phi^{-1}(1-n^{-1})\} , & \text{for } i = n . \end{cases}$$

Without loss of generality, assume n is even. Then

$$\sum_{1}^{n} g_i J(i,n)/G_n = 2n^{-1} \sum_{i=2}^{\frac{1}{2}n} g_i(\Phi^{-1}\{(i-1)/n\} + \frac{1}{2}n^{-1}\{\phi\{\Phi^{-1}((i-\theta)/n)\}\}^{-1})$$

$$+ 2g_n\phi\{\Phi^{-1}(1/n)\} .$$

By (16), for $1 < i \leqq \frac{1}{2}n$,

$$\phi\{\Phi^{-1}((i-\theta)/n)\}) \geqq \phi\{\Phi^{-1}((i-1)/n) \geqq \begin{cases} (i-1)\Phi^{-1}\{(i-1)/n\}/n, & 2 \leqq i \leqq kn,\ k < \frac{1}{2} \\ C(k), & kn < i \leqq \frac{1}{2}n . \end{cases}$$

Thus by Schwarz inequality,

$$|n^{-2} \sum_{2}^{\frac{1}{2}n} g_i/\phi\{\Phi^{-1}((i-\theta)/n)\}| \leqq n^{-\frac{1}{2}} G_n\{ \sum_{1}^{kn-1} \{i\Phi^{-1}(i/n)\}^{-2} + \sum_{kn+1}^{\frac{1}{2}n} C(k)^{-2}n^{-2}\}$$

which in turn is bounded by $C\{n\log(n)\}^{-1/2}$, in view of (19). Further, by (16) and (19), $\phi(\Phi^{-1}(1/n)) \sim O((\log n)/n)$, by (1), $g_n \sim m_n$ and with (32) and finally (22) we can argue that

$$2n^{-1}\sum_{2}^{\frac{1}{2}n} g_i\Phi^{-1}\{(i-1)/n\} \sim n^{-1}\sum_{1}^{n} g_i\Phi^{-1}\{i/(n+1)\} \sim n^{-1}\underline{m}'\underline{g} \sim M_nG_n \ .$$

These ensure that (7) holds.

4. Acknowledgements. We wish to thank the referees for their helpful comments and to thank one referee in particular whose very careful reading of the manuscript led to many improvements.

References

Balakrishnam, N. (1984). Approximating the Sum of Squares of Normal Scores. Appl. Statist., 33, 242-245.

David, H.A. (1981). Order Statistics. (2nd. Edition). John Wiley and Sons, Inc., New York.

De Wet, T. and Venter, J.H. (1972). Asymptotic Distributions of Certain Test Criteria of Normality. S. Afr. Statist. J., 6, 135-149.

Fotopoulos, S., Leslie, J.R. and Stephens, M.A. (1984). Errors in Approximations for Expected Normal Order Statistics with an Application to Goodness-of-Fit. Stanford Tech. Report No. 342.

Gregory, G.G. (1977). Functions of Order Statistics and Tests of Fit. South African Statist. J., 11, 99-118.

Hoeffding, W. (1953). On the Distribution of the Expected Values of the Order Statistics. Ann. Math. Statist., 24, 93-100.

Hogg, R.V. and Craig, A.T. (1970). Introduction to Mathematical Statistics. (3rd. Edition). MacMillan, New York.

Knopp, K. (1951). Theory and Application of Infinite Series. 2nd English Ed. Blackie and Sons Ltd., London and Glasgow.

Leslie, J.R. (1984). Asymptotic Properties and a New Approximation for Both the Covariance Matrix of Normal Order Statistics and its Inverse. Colloquia Math. Soc. J. Bolyai on Goodness of Fit. (to appear).

Renyi, A. (1970). Probability Theory. North-Holland, Amsterdam.

Ruben, H. (1956). On the sum of squares of normal scores. Biometrika, 43, 456-458.

Sarkadi, K. (1975). The Consistency of the Shapiro-Francia Test. Biometrika, 62, 455-450.

Sarkadi, K. (1977). Characterization and Testing for Normality. Colloquia Math. Soc. J. Bolyia, 21, 317-329.

Sarkadi, K. (1981). On the Consistency of some Goodness of Fit Tests. Proc. Sixth Conference Probab. Theory (Brasov, 1979), 195-204. Bucuresti, Ed. Acad. R.S. Romania.

Saw, J.G. and Chow, B. (1966). The Curve Through the Expected Values of Ordered Variates and the Sum of Squares of Normal Scores. Biometrika, 26, 252-255.

Shapiro, S.S. and Francia, R.S. (1972). An Approximation Analysis of Variance Test for Normality. J. Amer. Statist. Assoc., 67, 337, 215-216.

Shapiro, S.S. and Wilk, M.B. (1965). An Analysis of Variance test for Normality. Biometrika, 42, 591-611.

Stephens, M.A. (1975). Asymptotic Properties for Covariance Matrices of Order Statistics. Biometrika, 62, 1, 23-28.

Verrill, S.P. and Johnson, R.A. (1983). The Asymptotic Distributions of Censored Data Versions of the Shapiro-Wilk Test for Normality Statistics. Technical Report No. 702, Department of Statistics, University of Wisconsin.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER 411	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) Asymptotic Distribution Of The Shapiro-Wilk W For Testing For Normality		5. TYPE OF REPORT & PERIOD COVERED TECHNICAL REPORT
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) J.R. Leslie, M.A. Stephens, and S. Fotopoulos		8. CONTRACT OR GRANT NUMBER(s) N00014-86-K-0156
9. PERFORMING ORGANIZATION NAME AND ADDRESS Department of Statistics Stanford University Stanford, CA 94305		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS NR-042-267
11. CONTROLLING OFFICE NAME AND ADDRESS Office of Naval Research Statistics & Probability Program Code 1111		12. REPORT DATE December 22, 1988
		13. NUMBER OF PAGES 21
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)		15. SECURITY CLASS. (of this report) UNCLASSIFIED
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

APPROVED FOR PUBLIC RELEASE: DISTRIBUTION UNLIMITED

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Shapiro-Wilk statistic, goodness-of-fit, normal order scores, tests of normality.

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

Twenty years have elapsed since the Shapiro-Wilk statistic W for testing the normality of a sample first appeared. In that time a number of statistics which are close relatives of W have been found to have a common (known) asymptotic distribution. It was assumed therefore that W must have that asymptotic distribution. ~~We~~ show this to be the case and examine the norming constants that are used with all the statistics. In addition the consistency of the W-test is established.